

Optimal Execution for Uncertain Market Impact: Derivation and Characterization of a Continuous-Time Value Function

Kensuke Ishitani¹ and Takashi Kato²

¹School of Business Administration, Faculty of Urban Liberal Arts, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji, Tokyo 192-0397, Japan,
E-mail: ishitani-kensuke@tmu.ac.jp

²Division of Mathematical Science for Social Systems, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560-8531, Japan, E-mail: kato@sigmath.es.osaka-u.ac.jp

In this paper, we study an optimal execution problem in the case of uncertainty in market impact to derive a more realistic market model. Our model is a generalized version of that in [6], where a model of optimal execution with deterministic market impact was formulated. First, we construct a discrete-time model as a value function of an optimal execution problem. We express the market impact function as a product of a deterministic part (an increasing function with respect to the trader's execution volume) and a noise part (a positive random variable). Then, we derive a continuous-time model as a limit of a discrete-time value function. We find that the continuous-time value function is characterized by an optimal control problem with a Lévy process and investigate some of its properties, which are mathematical generalizations of the results in [6]. We also consider a typical example of the execution problem for a risk-neutral trader under log-linear/quadratic market impact with Gamma-distributed noise.

Keywords: market liquidity, optimal execution, uncertain market impact, Lévy process, viscosity solution, Hamilton–Jacobi–Bellman equation

1 Introduction

The optimal portfolio management problem is central in mathematical finance theory. There are various studies on this problem, and recently more realistic problems, such as liquidity problems, have attracted considerable attention. In this paper, we focus on market impact (MI), which is the effect of the investment behavior of traders on security prices. MI plays an important role in portfolio theory, and is also significant when we consider the case of an optimal execution problem, where a trader has a certain amount of security holdings (shares of a security held) and attempts to liquidate them before the time horizon. The

optimal execution problem with MI has been studied in several papers ([1], [2], [3], [4], [10] and references therein,) and in [6] such a problem is formulated mathematically.

It is often assumed that the MI function is deterministic. This assumption means that we can obtain information about MI in advance. However, in a real market it is difficult to estimate the effects of MI. Moreover, it often happens that a high concentration of unexpected orders will result in overfluctuation of the price. The Flash Crash in the United States stock market is a notable precedent of unusual thinning liquidity: On May 6th, 2010, the Dow Jones Industrial Average plunged by about 9%, only to recover the losses within minutes. Considering the uncertainty in MI, it is thus more realistic and meaningful to construct a mathematical model of random MI. Moazeni et al. [9] studied the uncertainty in MI caused by other institutions by compound Poisson processes, and then studied an optimization problem of expected proceeds of execution in a discrete-time setting. They considered the uncertainty in arrival times of large trades from other institutions; however, MI functions of decision makers themselves were given as deterministic linear functions so that the decision makers knew how their own execution affected the market price of the security (the coefficients of MI functions were regarded as “expected price depressions caused by trading assets at a unit rate”).

In this paper, we generalize the framework in [6], particularly considering a random MI function. We follow the approach in [6]: we construct a discrete-time model and take a limit to derive a continuous-time model and a corresponding value function. We assume that noise in the MI function in a discrete-time model is independent of both time and trading volume. The randomness of MI in the continuous-time model is described as a jump of a Lévy process. We study some properties of the continuous-time value function which are mathematical generalizations of the results in [6]. In particular, we find that the value function is characterized as a viscosity solution of the corresponding Hamilton-Jacobi-Bellman equation (abbreviated as HJB) when the MI function is sufficiently strong. We also perform a comparison with the case of deterministic MI and show that noise in MI makes a risk-neutral trader underestimate the MI cost. This means that a trader attempting to minimize the expected liquidation cost is not particularly sensitive to the uncertainty in MI. Moreover, we present generalizations of examples in [6] and investigate the effects of noise in MI on the optimal strategy of a trader by numerical experiments. We consider a risk-neutral trader execution problem with a log-linear/quadratic MI function whose noise is given by the Gamma distribution.

The rest of this paper is organized as follows. In Section 2, we present the mathematical formulation of our model. We set a discrete-time model of an optimal execution problem as our basic model and define the corresponding value function. In Section 3, we present our results, showing that the continuous-time value function is derived as a limit of the discrete-time one. Moreover, we investigate the continuity of the derived value function. Note that the results in this section are of the same form as those in [6]. In Section 4, we study the characterization of the value function as a viscosity solution of the corresponding HJB equation as a direct consequence of the result in [6]. In Section 5, we consider the case where the trader must sell all the shares of the security, which is referred to as a “sell-off condition.” We also study the optimization problem under the sell-off condition and show that the results in Section 4 in [6] also hold in our model. Section 6 treats the comparison between deterministic MI and random MI in a risk-neutral framework. In Section 7 we present some examples based on the proposed model. We conclude this paper in Section 8.

2 The Model

In this section, we present the details of the proposed model. Let (Ω, \mathcal{F}, P) be a complete probability space. $T > 0$ denotes a time horizon, and we assume $T = 1$ for brevity. We assume that the market consists of one risk-free asset (cash) and one risky asset (a security). The price of cash is always 1, which means that a risk-free rate is zero. The price of the security fluctuates according to a certain stochastic flow, and is influenced by sales performed by traders.

First, we consider a discrete-time model with a time interval $1/n$. We consider a single trader who has an endowment of $\Phi_0 > 0$ shares of a security. This trader liquidates the shares Φ_0 over a time interval $[0, 1]$ considering the effects of MI with noise. We assume that the trader sells shares at only times $0, 1/n, \dots, (n-1)/n$ for $n \in \mathbb{N} = \{1, 2, 3, \dots\}$.

For $l = 0, \dots, n$, we denote by S_l^n the price of the security at time l/n , and we also denote $X_l^n = \log S_l^n$. Let $s_0 > 0$ be an initial price (i.e., $S_0^n = s_0$) and $X_0^n = \log s_0$. If the trader sells an amount ψ_l^n at time l/n , the log price changes to $X_l^n - g_l^n(\psi_l^n)$, and by this execution (selling) the trader obtains an amount of cash $\psi_l^n S_l^n \exp(-g_l^n(\psi_l^n))$ as proceeds. Here, the random function

$$g_l^n(\psi, \omega) = c_l^n(\omega) g_n(\psi), \quad \psi \in [0, \Phi_0], \quad \omega \in \Omega$$

denotes MI with noise, which is given by the product of a positive random variable c_l^n and a deterministic function $g_n : [0, \Phi_0] \rightarrow [0, \infty)$. The function g_n is assumed to be non-decreasing, continuously differentiable and satisfying $g_n(0) = 0$. Moreover, we assume that $(c_l^n)_l$ is i.i.d., and therefore noise in MI is time-homogeneous. Note that if c_l^n is a constant (i.e., $c_l^n \equiv c$ for some $c > 0$), then this setting is the same as in [6].

After trading at time l/n , X_{l+1}^n and S_{l+1}^n are given by

$$X_{l+1}^n = Y\left(\frac{l+1}{n}; \frac{l}{n}, X_l^n - g_l^n(\psi_l^n)\right), \quad S_{l+1}^n = e^{X_{l+1}^n}, \quad (2.1)$$

where $Y(t; r, x)$ is the solution of the following stochastic differential equation (SDE) on the filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t^B)_t, P)$

$$\begin{cases} dY(t; r, x) = \sigma(Y(t; r, x))dB_t + b(Y(t; r, x))dt, & t \geq r, \\ Y(r; r, x) = x, \end{cases}$$

where $(B_t)_{0 \leq t \leq 1}$ is standard one-dimensional Brownian motion which is independent of $(c_l^n)_l$, $(\mathcal{F}_t^B)_t$ is its Brownian filtration, and $b, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ are Borel functions. We assume that b and σ are bounded and Lipschitz continuous. Then, for each $r \geq 0$ and $x \in \mathbb{R}$, there exists a unique solution.

At the end of the time interval $[0, 1]$, the trader has an amount of cash W_n^n and an amount of the security φ_n^n , where

$$W_{l+1}^n = W_l^n + \psi_l^n S_l^n e^{-g_l^n(\psi_l^n)}, \quad \varphi_{l+1}^n = \varphi_l^n - \psi_l^n \quad (2.2)$$

for $l = 0, \dots, n-1$ and $W_0^n = 0$, $\varphi_0^n = \Phi_0$. We say that an execution strategy $(\psi_l^n)_{l=0}^{n-1}$ is admissible if $(\psi_l^n)_l \in \mathcal{A}_k^n(\Phi_0)$ holds, where $\mathcal{A}_k^n(\varphi)$ is the set of strategies $(\psi_l^n)_{l=0}^{k-1}$ such that ψ_l^n

is $\mathcal{F}_l^n = \sigma\{(B_t)_{t \leq l/n}, c_0^n, \dots, c_{l-1}^n\}$ -measurable, $\psi_l^n \geq 0$ for each $l = 0, \dots, k-1$ and $\sum_{l=0}^{k-1} \psi_l^n \leq \varphi$ almost surely.

Then, the investor's problem is to choose an admissible strategy to maximize the expected utility $E[u(W_n^n, \varphi_n^n, S_n^n)]$, where $u \in \mathcal{C}$ is the utility function employed by the investor and \mathcal{C} is the set of non-decreasing, non-negative and continuous functions on $D = \mathbb{R} \times [0, \Phi_0] \times [0, \infty)$ such that

$$u(w, \varphi, s) \leq C_u(1 + |w|^{m_u} + s^{m_u}), \quad (w, \varphi, s) \in D \quad (2.3)$$

for some constants $C_u, m_u > 0$.

For $k = 1, \dots, n$, $(w, \varphi, s) \in D$ and $u \in \mathcal{C}$, we define the (discrete-time) value function $V_k^n(w, \varphi, s; u)$ by

$$V_k^n(w, \varphi, s; u) = \sup_{(\psi_l^n)_{l=0}^{k-1} \in \mathcal{A}_k^n(\varphi)} E[u(W_k^n, \varphi_k^n, S_k^n)]$$

subject to (2.1) and (2.2) for $l = 0, \dots, k-1$ and $(W_0^n, \varphi_0^n, S_0^n) = (w, \varphi, s)$ (for $s = 0$, we set $S_l^n \equiv 0$). We denote such a triplet of processes $(W_l^n, \varphi_l^n, S_l^n)_{l=0}^k$ by $\Xi_k^n(w, \varphi, s; (\psi_l^n)_l)$, and denote $V_0^n(w, \varphi, s; u) = u(w, \varphi, s)$. Then, this problem is equivalent to $V_n^n(0, \Phi_0, s_0; u)$. We consider the limit of the value function $V_k^n(w, \varphi, s; u)$ as $n \rightarrow \infty$.

Let $h : [0, \infty) \rightarrow [0, \infty)$ be a non-decreasing continuous function. We introduce the following condition for $g_n(\psi)$.

$$[\mathbf{A}] \quad \lim_{n \rightarrow \infty} \sup_{\psi \in [0, \Phi_0]} \left| \frac{d}{d\psi} g_n(\psi) - h(n\psi) \right| = 0.$$

Moreover, we assume the following conditions for $(c_l^n)_l$.

[B1] Define $\gamma_n = \text{essinf}_{\omega} c_l^n(\omega)$. For any $n \in \mathbb{N}$, it holds that $\gamma_n > 0$. In addition,

$$\frac{h(x/\gamma_n)}{n} \rightarrow 0, \quad n \rightarrow \infty \quad (2.4)$$

holds for $x \geq 0$.

[B2] Let μ_n be the distribution of $\frac{c_0^n + \dots + c_{n-1}^n}{n}$. Then, μ_n has a weak limit μ as $n \rightarrow \infty$.

[B3] There is a sequence of infinitely divisible distributions $(p_n)_n$ on \mathbb{R} such that $\mu_n = \mu * p_n$ and either

$$[\mathbf{B3-a}] \quad \int_{\mathbb{R}} x^2 p_n(dx) = O(1/n^3) \text{ as } n \rightarrow \infty$$

or

[B3-b] There is a sequence $(K_n)_n \subset (0, \infty)$ such that $K_n = O(1/n)$, $p_n((-\infty, -K_n)) = 0$ (or $p_n((K_n, \infty)) = 0$) and $\int_{\mathbb{R}} x p_n(dx) = O(1/n)$ as $n \rightarrow \infty$, where O denotes the order notation (Landau's symbol).

Let us discuss condition [B1]. First, note that γ_n is independent of l , because c_l^n , $l = 0, 1, 2, \dots$ are identically distributed. Next, we study when the convergence (2.4) holds. Since h is non-decreasing, we see that

$$\frac{h(x/\gamma_n)}{n} \leq \frac{h(\infty)}{n}, \quad n \in \mathbb{N},$$

where $h(\infty) = \lim_{\zeta \rightarrow \infty} h(\zeta) \in [0, \infty]$ (which is well-defined by virtue of the monotonicity of h). This inequality tells us that (2.4) is always fulfilled whenever $h(\infty) < \infty$. In the case of $h(\infty) = \infty$, we have the following example:

$$h(\zeta) = \alpha \zeta^p, \quad \gamma_n = \frac{1}{n^{1/p-\delta}} \quad (p, \delta > 0, \delta \leq 1/p). \quad (2.5)$$

We can actually confirm (2.4) by observing

$$\frac{h(x/\gamma_n)}{n} = \frac{\alpha x^p}{n^{p\delta}} \longrightarrow 0, \quad n \rightarrow \infty.$$

Here, the condition [B2] holds only when $\limsup_n \gamma_n < \infty$. Indeed, under [B2], we easily see that the support of the distribution μ is included in the interval $[\limsup_n \gamma_n, \infty)$. Note that γ_n of (2.5) satisfies $\limsup_n \gamma_n \leq 1$ because of the relation $\delta \leq 1/p$.

Since μ is an infinitely divisible distribution, there is some Lévy process (subordinator) $(L_t)_{0 \leq t \leq 1}$ on a certain probability space such that L_1 is distributed by μ . Without loss of generality, we may assume that $(L_t)_t$ and $(B_t)_t$ are defined on the same filtered space. Since $(c_l^n)_l$ is independent of $(B_t)_t$, we may also assume that $(L_t)_t$ is independent of $(B_t)_t$. Let ν be the Lévy measure of $(L_t)_t$. We assume the following moment condition for ν .

$$[\text{C}] \quad \int_{(0, \infty)} (z + z^2) \nu(dz) < \infty.$$

Now, we define the function which gives the limit of the discrete-time value function. For $t \in [0, 1]$ and $\varphi \in [0, \Phi_0]$ we denote by $\mathcal{A}_t(\varphi)$ the set of $(\mathcal{F}_r)_{0 \leq r \leq t}$ -adapted and càglàd processes (i.e., left-continuous and having a right limit at each point) $(\zeta_r)_{0 \leq r \leq t}$ such that $\zeta_r \geq 0$ for each

$r \in [0, t]$, $\int_0^t \zeta_r dr \leq \varphi$ almost surely and

$$\sup_{(r, \omega) \in [0, t] \times \Omega} \zeta_r(\omega) < \infty, \quad (2.6)$$

where $\mathcal{F}_r = \sigma\{B_v, L_v; v \leq r\} \vee \{\text{Null Sets}\}$. Here, the supremum in (2.6) is taken over all values in $[0, t] \times \Omega$. Note that we may replace \sup in (2.6) with esssup .

For $t \in [0, 1]$, $(w, \varphi, s) \in D$ and $u \in \mathcal{C}$, we define $V_t(w, \varphi, s; u)$ by

$$V_t(w, \varphi, s; u) = \sup_{(\zeta_r)_{r \in \mathcal{A}_t(\varphi)}} \mathbb{E}[u(W_t, \varphi_t, S_t)] \quad (2.7)$$

subject to

$$\begin{aligned} dW_r &= \zeta_r S_r dr, \\ d\varphi_r &= -\zeta_r dr, \\ dX_r &= \sigma(X_r)dB_r + b(X_r)dr - g(\zeta_r)dL_r, \quad S_r = \exp(X_r) \end{aligned}$$

and $(W_0, \varphi_0, S_0) = (w, \varphi, s)$, where $\hat{\sigma}(s) = s\sigma(\log s)$, $\hat{b}(s) = s\left\{b(\log s) + \frac{1}{2}\sigma(\log s)^2\right\}$ for $s > 0$ ($\hat{\sigma}(0) = \hat{b}(0) = 0$) and $g(\zeta) = \int_0^\zeta h(\zeta')d\zeta'$. We denote such a triplet of processes $(W_r, \varphi_r, S_r)_{0 \leq r \leq t}$ by $\Xi_t(w, \varphi, s; (\zeta_r)_r)$. Note that $V_0(w, \varphi, s; u) = u(w, \varphi, s)$. We call $V_t(w, \varphi, s; u)$ a continuous-time value function. Also note that $V_t(w, \varphi, s; u) < \infty$ for each $t \in [0, 1]$ and $(w, \varphi, s) \in D$.

3 Properties of Value Functions

First, we give the convergence theorem for value functions.

Theorem 1. *For each $(w, \varphi, s) \in D$, $t \in [0, 1]$ and $u \in \mathcal{C}$ it holds that*

$$\lim_{n \rightarrow \infty} V_{[nt]}^n(w, \varphi, s; u) = V_t(w, \varphi, s; u),$$

where $[nt]$ is the greatest integer $\leq nt$.

According to this theorem, a discrete-time value function converges to $V_t(w, \varphi, s; u)$ by shortening the time intervals of execution. This implies that we can regard $V_t(w, \varphi, s; u)$ as the value function of the continuous-time model of an optimal execution problem with random MI. This result is almost the same as in [6], with the exception that the term of MI is given as an increment $g(\zeta_r)dL_r$. Let

$$L_t = \gamma t + \int_0^t \int_{(0, \infty)} z N(dr, dz)$$

be the Lévy decomposition of $(L_t)_t$, where $\gamma \geq 0$ and $N(\cdot, \cdot)$ is a Poisson random measure. Then, $g(\zeta_r)dL_r$ can be divided into two terms as follows:

$$g(\zeta_{r-})dL_r = \gamma g(\zeta_r)dr + g(\zeta_{r-}) \int_{(0, \infty)} z N(dr, dz).$$

The last term on the right-hand side indicates the effect of noise in MI. This means that noise in MI appears as a jump of a Lévy process. Using the above representation and Itô's formula, we see that when $s > 0$, the process $(S_r)_r$ satisfies

$$dS_r = \hat{\sigma}(S_r)dB_r + \hat{b}(S_r)dr - \left\{ \gamma g(\zeta_r)S_r dr + S_{r-} \int_{(0, \infty)} (1 - e^{-g(\zeta_{r-})z}) N(dz, dr) \right\}.$$

Regarding the continuity of the continuous-time value function, we have the following theorem.

Theorem 2. *Let $u \in \mathcal{C}$.*

- (i) *If $h(\infty) = \infty$, then $V_t(w, \varphi, s; u)$ is continuous in $(t, w, \varphi, s) \in [0, 1] \times D$.*
- (ii) *If $h(\infty) < \infty$, then $V_t(w, \varphi, s; u)$ is continuous in $(t, w, \varphi, s) \in (0, 1] \times D$ and $V_t(w, \varphi, s; u)$ converges to $Ju(w, \varphi, s)$ uniformly on any compact subset of D as $t \downarrow 0$, where $Ju(w, \varphi, s)$ is given as*

$$\begin{cases} \sup_{\psi \in [0, \varphi]} u\left(w + \frac{1 - e^{-\gamma h(\infty)\psi}}{\gamma h(\infty)}s, \varphi - \psi, se^{-\gamma h(\infty)\psi}\right) & (\gamma h(\infty) > 0), \\ \sup_{\psi \in [0, \varphi]} u(w + \psi s, \varphi - \psi, s) & (\gamma h(\infty) = 0). \end{cases}$$

This is also quite similar to the result in [6], whereby continuities in w, φ and s of the value function are always guaranteed, but continuity in t at the origin depends on the state of the function h at infinity. When $h(\infty) = \infty$, MI for large sales is sufficiently strong ($g(\zeta)$ diverges rapidly with $\zeta \rightarrow \infty$) to make a trader avoid instant liquidation: an optimal policy is “no trading” in infinitesimal time, and thus V_t converges to u as $t \downarrow 0$. When $h(\infty) < \infty$, the value function is not always continuous at $t = 0$ and has the right limit $Ju(w, \varphi, s)$. In this case, MI for large sales is not particularly strong ($g(\zeta)$ still diverges, but the divergence speed is low) and there is room for liquidation within infinitesimal time. The function $Ju(w, \varphi, s)$ corresponds to the utility of the liquidation of the trader, who sells part of the shares of a security ψ by dividing it infinitely within an infinitely short time (sufficiently short that the fluctuation in the price of the security can be ignored) and obtains an amount $\varphi - \psi$, that is,

$$\zeta_r^\delta = \frac{\psi}{\delta} 1_{[0, \delta]}(r), \quad r \in [0, t] \quad (\delta \downarrow 0). \quad (3.1)$$

We pay attention to the fact that the jump part $g(\zeta_{r-}) \int_{(0, \infty)} z N(dr, dz)$ does not change the result. Note that if $\gamma = 0$ and $h(\infty) < \infty$, then the effect of MI disappears in $Ju(w, \varphi, s)$. This situation may occur even if $E[c_k^n] \geq \varepsilon_0$ (or $E[L_1] \geq \varepsilon_0$) for some $\varepsilon_0 > 0$.

For the proofs of Theorems 1–2, please refer to [5].

4 Characterization of the Value Function as a Viscosity Solution of a Corresponding HJB

As presented in [6], we can characterize our value function as a viscosity solution of a corresponding HJB when $h(\infty) = \infty$. First, we present the Bellman principle (dynamic programming principle). Let us define $Q_t : \mathcal{C} \rightarrow \mathcal{C}$ by $Q_t u(w, \varphi, s) = V_t(w, \varphi, s; u)$. Then, we can easily confirm that Q_t is well-defined as a nonlinear operator. The same proof as that for Theorem 3 in [6] gives the following proposition.

Proposition 1. *For each $r, t \in [0, 1]$ with $t + r \leq 1$, $(w, \varphi, s) \in D$ and $u \in \mathcal{C}$, it holds that $Q_{t+r}u(w, \varphi, s) = Q_t Q_r u(w, \varphi, s)$.*

Note that this proposition is also required to prove Theorem 2. At this stage, we introduce the HJB corresponding to our value function. We start by formal calculation to derive the

HJB and subsequently show that we can apply Theorems 4–5 in [6] directly to our case, in spite of the fact that our value function is a generalization of that in [6].

We begin by fixing $t \in (0, 1]$ and letting $h > 0$ be such that $t - h \geq 0$. Then, Proposition 1 leads us to

$$V_t(w, \varphi, s; u) = Q_h Q_{t-h} u(w, \varphi, s; u) = \sup_{(\zeta_r)_{r \in \mathcal{A}_h(\varphi)}} \mathbb{E}[V_{t-h}(W_h, \varphi_h, S_h; u)].$$

If we restrict the admissible strategies to constant ones and assume the smoothness of $V_t(w, \varphi, s; u)$, we obtain

$$\begin{aligned} 0 &= \sup_{\zeta \geq 0} \mathbb{E}[V_{t-h}(W_h, \varphi_h, S_h; u)] - V_t(w, \varphi, s; u) \\ &= \sup_{\zeta \geq 0} \mathbb{E}\left[\int_0^h \left(-\frac{\partial}{\partial t} + \mathcal{L}^\zeta\right) V_r(W_r, \varphi_r, S_r; u) dr\right] \end{aligned}$$

to arrive at

$$\frac{1}{h} \sup_{\zeta \geq 0} \mathbb{E}\left[\int_0^h \left(-\frac{\partial}{\partial t} + \mathcal{L}^\zeta\right) V_r(W_r, \varphi_r, S_r; u) dr\right] = 0 \quad (4.1)$$

by virtue of Itô's formula, where

$$\begin{aligned} \mathcal{L}^\zeta v(t, w, \varphi, s) &= \frac{1}{2} \hat{\sigma}(s)^2 \frac{\partial^2}{\partial s^2} v(t, w, \varphi, s) + \hat{b}(s) \frac{\partial}{\partial s} v(t, w, \varphi, s) \\ &\quad + \zeta \left(s \frac{\partial}{\partial w} v(t, w, \varphi, s) - \frac{\partial}{\partial \varphi} v(t, w, \varphi, s) \right) - \hat{g}(\zeta) s \frac{\partial}{\partial s} v(t, w, \varphi, s), \\ \hat{g}(\zeta) &= \gamma g(\zeta) + \int_{(0, \infty)} (1 - e^{-g(\zeta)z}) \nu(dz). \end{aligned}$$

We remark that \mathcal{L}^ζ is of exactly the same form as in [6]. Letting $h \rightarrow 0$ in (4.1), we obtain

$$\frac{\partial}{\partial t} V_r(W_r, \varphi_r, S_r; u) - \sup_{\zeta \geq 0} \mathcal{L}^\zeta V_r(W_r, \varphi_r, S_r; u) = 0.$$

Note that this calculation is nothing but intuitive and formal. However, we can justify the following theorem.

Theorem 3. *Assume that h is strictly increasing and $h(\infty) = \infty$. Moreover, assume*

$$\liminf_{\varepsilon \downarrow 0} \frac{V_t(w, \varphi, s + \varepsilon; u) - V_t(w, \varphi, s; u)}{\varepsilon} > 0 \quad (4.2)$$

for any $t \in (0, 1]$ and $(w, \varphi, s) \in U$. Then, $V_t(w, \varphi, s; u)$ is a viscosity solution of

$$\frac{\partial}{\partial t} v(t, w, \varphi, s) - \sup_{\zeta \geq 0} \mathcal{L}^\zeta v(t, w, \varphi, s) = 0, \quad (t, w, \varphi, s) \in (0, 1] \times U, \quad (4.3)$$

where $U = \hat{D} \setminus \partial \hat{D}$ and $\hat{D} = \mathbb{R} \times [0, \infty) \times [0, \infty)$.

Proof. We consider the strategy-restricted version of the value function

$$V_t^L(w, \varphi, s; u) = \sup_{(\zeta_r)_{r \leq t} \in \mathcal{A}_t^L(\varphi)} \mathbb{E}[u(W_t, \varphi_t, S_t)]$$

for $L > 0$, where $\mathcal{A}_t^L(\varphi) = \{(\zeta_r)_{0 \leq r \leq t} \in \mathcal{A}_t(\varphi) ; \sup_{r, \omega} |\zeta_r(\omega)| \leq L\}$. Then, the standard argument (see Section 5.4 in [11]) suggests that $V_t^L(w, \varphi, s; u)$ is a viscosity solution of

$$\frac{\partial}{\partial t} v(t, w, \varphi, s) - \sup_{0 \leq \zeta \leq L} \mathcal{L}^\zeta v(t, w, \varphi, s) = 0, \quad (t, w, \varphi, s) \in (0, 1] \times U$$

because of the compactness of the control region $[0, L]$. Here, we can define $\hat{h}(\zeta)$ as the derivative of \hat{g} by

$$\hat{h}(\zeta) = \left(\gamma + \int_{(0, \infty)} z e^{-g(\zeta)z} \nu(dz) \right) h(\zeta).$$

Obviously, it holds that $\hat{h}(\infty) = \infty$. Hence, we can apply the same argument as in Section 7.6 in [6] to obtain the assertion by letting $L \rightarrow \infty$ and using the stability arguments for viscosity solutions. \blacksquare

Since the condition $\lim_{\zeta \rightarrow \infty} (h(\zeta)/\zeta) > 0$ implies $\lim_{\zeta \rightarrow \infty} (\hat{h}(\zeta)/\zeta) > 0$, we can apply Theorem 5 in [6] to arrive at the following uniqueness theorem.

Theorem 4. *Assume that $\hat{\sigma}$ and \hat{b} are both Lipschitz continuous. Assume the hypotheses in Theorem 3 and that $\liminf_{\zeta \rightarrow \infty} (h(\zeta)/\zeta) > 0$. If a polynomial growth function $v : [0, 1] \times \hat{D} \rightarrow \mathbb{R}$ is a viscosity solution of (4.3) and satisfies the boundary conditions*

$$\begin{aligned} v(0, w, \varphi, s) &= u(w, \varphi, s), \quad (w, \varphi, s) \in \hat{D}, \\ v(t, w, 0, s) &= \mathbb{E}[u(w, 0, Z(t; 0, s))], \quad (t, w, s) \in [0, 1] \times \mathbb{R} \times [0, \infty), \\ v(t, w, \varphi, 0) &= u(w, \varphi, 0), \quad (t, w, \varphi) \in [0, 1] \times \mathbb{R} \times [0, \infty), \end{aligned} \tag{4.4}$$

then $V_t(w, \varphi, s; u) = v(t, w, \varphi, s)$, where

$$Z(t; r, s) = \exp(Y(t; r, \log s)) \quad (s > 0), \quad 0 \quad (s = 0). \tag{4.5}$$

5 Sell-Off Condition

In this section, we consider the optimal execution problem under the “sell-off condition”, which was introduced in [6]. A trader has a certain quantity of shares of a security at the initial time, and must liquidate all of them by the time horizon. Then, the spaces of admissible strategies are reduced to

$$\begin{aligned} \mathcal{A}_k^{n, \text{SO}}(\varphi) &= \left\{ (\psi_l^n)_l \in \mathcal{A}_k^n(\varphi) ; \sum_{l=0}^{k-1} \psi_l^n = \varphi \right\}, \\ \mathcal{A}_t^{\text{SO}}(\varphi) &= \left\{ (\zeta_r)_r \in \mathcal{A}_t(\varphi) ; \int_0^t \zeta_r dr = \varphi \right\}. \end{aligned}$$

Now, we define value functions with the sell-off condition by

$$\begin{aligned} V_k^{n,\text{SO}}(w, \varphi, s; U) &= \sup_{(\psi_l^n)_{l \in \mathcal{A}_k^{n,\text{SO}}(\varphi)}} \mathbb{E}[U(W_k^n)], \\ V_t^{\text{SO}}(w, \varphi, s; U) &= \sup_{(\zeta_r)_{r \in \mathcal{A}_t^{\text{SO}}(\varphi)}} \mathbb{E}[U(W_t)] \end{aligned}$$

for a continuous, non-decreasing and polynomial growth function $U : \mathbb{R} \rightarrow \mathbb{R}$, where $(W_l^n, \varphi_l^n, S_l^n)_{l=0}^k = \Xi_k^n(w, \varphi, s; (\psi_l^n)_l)$ and $(W_r, \varphi_r, S_r)_{0 \leq r \leq t} = \Xi_t(w, \varphi, s; (\zeta_r)_r)$.

The following theorem is analogous to Theorem 7 in [6].

Theorem 5. $V_t^{\text{SO}}(w, \varphi, s; U) = V_t(w, \varphi, s; u)$, where $u(w, \varphi, s) = U(w)$.

Proof. The relation $V_t^{\text{SO}}(w, \varphi, s; U) \leq V_t(w, \varphi, s; u)$ is trivial, so we show only the assertion $V_t^{\text{SO}}(w, \varphi, s; U) \geq V_t(w, \varphi, s; u)$. Take any $(\zeta_r)_r \in \mathcal{A}_t(\varphi)$ and let $(W_r, \varphi_r, S_r)_r = \Xi_1(w, \varphi, s; (\zeta_r)_r)$. Also, take any $\delta \in (0, t)$. We define an execution strategy $(\zeta_r^\delta)_r \in \mathcal{A}_t^{\text{SO}}(\varphi)$ by $\zeta_r^\delta = \zeta_r$ ($r \in [0, t - \delta]$), $\varphi_{t-\delta}/\delta$ ($r \in (t - \delta, t]$). Let $(W_r^\delta, \varphi_r^\delta, S_r^\delta)_r = \Xi_t(w, \varphi, s; (\zeta_r^\delta)_r)$. Then, we have $W_{t-\delta} = W_{t-\delta}^\delta \leq W_t^\delta$, arriving at $\mathbb{E}[U(W_{t-\delta})] \leq \mathbb{E}[U(W_t^\delta)] \leq V_t^{\text{SO}}(w, \varphi, s; U)$. Letting $\delta \downarrow 0$, we obtain $\mathbb{E}[U(W_t)] \leq V_t^{\text{SO}}(w, \varphi, s; U)$ by using the monotone convergence theorem. Since $(\zeta_r)_r \in \mathcal{A}_t(\varphi)$ is arbitrary, we obtain the assertion. \blacksquare

By Theorem 5, we see that the sell-off condition $\int_0^t \zeta_r dr = \varphi$ does not introduce changes in the (value of the) value function in a continuous-time model. Thus, although the value function in a discrete-time model may depend on whether the sell-off condition is imposed, in the continuous-time model this condition is irrelevant.

The following is also similar to Theorem 7 in [6], which is a version of Theorem 1 with the sell-off condition.

Theorem 6. For any $(w, \varphi, s) \in D$,

$$\lim_{n \rightarrow \infty} V_{[nt]}^{n,\text{SO}}(w, \varphi, s; U) = V_t^{\text{SO}}(w, \varphi, s; U) \quad (= V_t(w, \varphi, s; U)).$$

Proof. We may assume $t > 0$. Take any $\delta \in (0, t)$ and let $n > 1/\delta$. Then, each strategy in $\mathcal{A}_{[n(t-\delta)]}^n(\varphi)$ can always be extended to the one in $\mathcal{A}_{[nt]}^{n,\text{SO}}(\varphi)$ by liquidating all remaining inventory in the last period. Thus, we see that for $n > 1/\delta$

$$V_{[n(t-\delta)]}^n(w, \varphi, s; u) \leq V_{[nt]}^{n,\text{SO}}(w, \varphi, s; U) \leq V_{[nt]}^n(w, \varphi, s; u). \quad (5.1)$$

By Theorem 1, we obtain

$$\lim_{n \rightarrow \infty} V_{[n(t-\delta)]}^n(w, \varphi, s; u) = V_{t-\delta}(w, \varphi, s; u), \quad \lim_{n \rightarrow \infty} V_{[nt]}^n(w, \varphi, s; u) = V_t(w, \varphi, s; u). \quad (5.2)$$

By (5.1), (5.2), and Theorem 2, we obtain the assertion. \blacksquare

Analogously to Theorem 8 in [6], a result similar to Theorem 3 in [8] holds when $g(\zeta)$ is linear:

Theorem 7. Assume $g(\zeta) = \alpha_0 \zeta$ for $\alpha_0 > 0$.

(i) $V_t^{\text{SO}}(w, \varphi, s; U) = \bar{V}_t^\varphi \left(w + \frac{1 - e^{-\gamma\alpha_0\varphi}}{\gamma\alpha_0} s, e^{-\gamma\alpha_0\varphi} s; U \right)$, where

$$\begin{aligned} \bar{V}_t^\varphi(\bar{w}, \bar{s}; U) &= \sup_{(\bar{\varphi}_r)_r \in \bar{\mathcal{A}}_t(\varphi)} \mathbb{E}[U(\bar{W}_t)] \\ \text{s.t.} \quad d\bar{S}_r &= e^{-\gamma\alpha_0\bar{\varphi}_r} \hat{b}(\bar{S}_r e^{\gamma\alpha_0\bar{\varphi}_r}) dr + e^{-\gamma\alpha_0\bar{\varphi}_r} \hat{\sigma}(\bar{S}_r e^{\gamma\alpha_0\bar{\varphi}_r}) dB_r - \bar{S}_r dG_r, \\ d\bar{W}_r &= \frac{e^{\gamma\alpha_0\bar{\varphi}_r} - 1}{\gamma\alpha_0} d\bar{S}_r, \\ \bar{S}_0 &= \bar{s}, \quad \bar{W}_0 = \bar{w} \end{aligned}$$

and

$$\bar{\mathcal{A}}_t(\varphi) = \left\{ \left(\varphi - \int_0^r \zeta_v dv \right)_{0 \leq r \leq t} ; (\zeta_r)_{0 \leq r \leq t} \in \mathcal{A}_t^{\text{SO}}(\varphi) \right\},$$

where

$$G_r = \int_0^r \int_{(0, \infty)} (1 - e^{-\alpha_0 \zeta_s - z}) N(ds, dz).$$

(ii) If U is concave and $\hat{b}(s) \leq 0$ for $s \geq 0$, then

$$V_t^{\text{SO}}(w, \varphi, s; U) = U \left(w + \frac{1 - e^{-\gamma\alpha_0\varphi}}{\gamma\alpha_0} s \right). \quad (5.3)$$

Proof. We can easily confirm assertion (i) by applying Itô's formula to \bar{S}_r and \bar{W}_r . By a similar argument to that in Section 7.9 in [6], we obtain

$$\mathbb{E}[U(\bar{W}_t)] \leq U \left(\bar{w} + \int_0^t \mathbb{E} \left[\frac{1 - e^{-\gamma\alpha_0\bar{\varphi}_r}}{\gamma\alpha_0} \hat{b}(\bar{S}_r e^{\gamma\alpha_0\bar{\varphi}_r}) - \int_{(0, \infty)} \frac{e^{\gamma\alpha_0\bar{\varphi}_r} - 1}{\gamma\alpha_0} \bar{S}_r (1 - e^{-\alpha_0 \zeta_r - z}) \nu(dz) \right] dr \right)$$

for any $(\bar{\varphi}_r)_r \in \bar{\mathcal{A}}_t(\varphi)$ by virtue of the Jensen inequality. Since \hat{b} is non-positive, the function U is non-decreasing, and the terms

$$1 - e^{-\gamma\alpha_0\bar{\varphi}_r}, \quad e^{\gamma\alpha_0\bar{\varphi}_r} - 1, \quad 1 - e^{-\alpha_0 \zeta_r - z}$$

are all non-negative, we see that $\mathbb{E}[U(\bar{W}_t)] \leq U(\bar{w})$ for any $(\bar{\varphi}_r)_r \in \bar{\mathcal{A}}_t(\varphi)$, which implies $\bar{V}_t^\varphi(\bar{w}, \bar{s}) \leq U(\bar{w})$. The opposite inequality $\bar{V}_t^\varphi(\bar{w}, \bar{s}) \geq U(\bar{w})$ is obtained, similarly to the result in Section 7.9 in [6] and the proof of Proposition 11 in [5]. \blacksquare

We note that the assertion (ii) is the same as Theorem 3 in [8], and in this case we can obtain the explicit form of the value function. The right-hand side of (5.3) is equal to $Ju(w, \varphi, s)$ for $u(w, \varphi, s) = U(w)$ and the nearly optimal strategy for $V_t^{\text{SO}}(w, \varphi, s; U) = V_t(w, \varphi, s; u)$ is given by (3.1). This implies that whenever considering a linear MI function, a risk-averse (or risk-neutral) trader's optimal strategy of liquidating a security that has a negative risk-adjusted drift is nearly the same as block liquidation (i.e. selling all shares at once) at the initial time.

6 Effect of Uncertainty in MI in the Risk-Neutral Framework

The purpose of this section is to investigate how the noise in the MI function affects the trader. Particularly, we focus on the case where the trader is risk-neutral, i.e. $u(w, \varphi, s) = u_{\text{RN}}(w, \varphi, s) = w$.

First, we prepare a value function of the execution problem with a deterministic MI function to perform a comparison with the case of random MI. Let $g_l^n(\psi)$ be as in Section 3, and set $\bar{g}_n(\psi) \equiv \mathbb{E}[g_l^n(\psi)] = \mathbb{E}[c_l^n]g_n(\psi)$. Note that when $\mathbb{E}[c_l^n] = 1$, the function \bar{g}_n is equivalent to g_n . We denote the discrete-time value function by $\bar{V}_k^n(w, \varphi, s; u)$ with an MI function \bar{g}_n . Theorem 1 in [6] implies that under [A] the function $\bar{V}_{[nt]}^n(w, \varphi, s; u)$ converges to the continuous-time value function $\bar{V}_t(w, \varphi, s; u)$, which becomes the same as in (2.7) by replacing $g(\zeta)$ and L_t with $\tilde{\gamma}g(\zeta)$ and t , i.e., the SDE for $(X_r)_r$ is given as

$$dX_r = \sigma(X_r)dB_r + b(X_r)dr - \tilde{\gamma}g(\zeta_r)dr,$$

where $\tilde{\gamma} = \mathbb{E}[L_1]$.

Theorem 8. $V_k^n(w, \varphi, s; u_{\text{RN}}) \geq \bar{V}_k^n(w, \varphi, s; u_{\text{RN}})$.

Proof. Take any $(\psi_l^n)_l \in \mathcal{A}_k^n(\varphi)$ and let $(W_l^n, \varphi_l^n, S_l^n)_l = \Xi_k^n(w, \varphi, s; (\psi_l^n)_l)$ be the triplet for $\bar{V}_k^n(w, \varphi, s; u_{\text{RN}})$. Then, the Jensen inequality implies

$$\begin{aligned} \mathbb{E}[W_k^n] &= w + \sum_{l=0}^{k-1} \mathbb{E}[\psi_l^n S_l^n \exp(-\bar{g}_n(\psi_l^n))] \\ &= w + \sum_{l=0}^{k-1} \mathbb{E}[\psi_l^n S_l^n \exp(-\mathbb{E}[c_l^n | \mathcal{F}_l^n] g_n(\psi_l^n))] \\ &\leq w + \sum_{l=0}^{k-1} \mathbb{E}[\psi_l^n S_l^n \mathbb{E}[\exp(-c_l^n g_n(\psi_l^n)) | \mathcal{F}_l^n]] \leq V_k^n(w, \varphi, s; u_{\text{RN}}). \end{aligned}$$

Since $(\psi_l^n)_l$ is arbitrary, we obtain the assertion. ■

The above proposition immediately leads us to

$$V_t(w, \varphi, s; u_{\text{RN}}) \geq \bar{V}_t(w, \varphi, s; u_{\text{RN}}). \quad (6.1)$$

This inequality shows that noise in MI is welcome since it decreases the liquidation cost for a risk-neutral trader.

For instance, we consider the situation where the trader estimates the MI function from historical data and tries to minimize the expected liquidation cost. Then, a higher sensitivity of the trader to the volatility risk of MI results in a lower estimate for the expected proceeds of the liquidation. This implies that taking into consideration the uncertainty in MI makes the trader prone to underestimating the liquidation cost. Thus, as long as the trader's target is the expected cost, considering the uncertainty in MI is not an incentive for being conservative with respect to the unpredictable liquidity risk. In Section 7, we present the results of numerical experiments conducted in order to simulate above phenomenon.

7 Examples

In this section, we show two examples of our model, both of which are generalizations of the ones in [6].

Motivated by the Black-Scholes type market model, we take $b(x) \equiv -\mu$ and $\sigma(x) \equiv \sigma$ for some constants $\mu, \sigma \geq 0$ and assume $\tilde{\mu} = \mu - \sigma^2/2 > 0$. We also assume that a trader has a risk-neutral utility function $u(w, \varphi, s) = u_{\text{RN}}(w) = w$. In this case, if there is no MI, then a risk-neutral trader is afraid of decreasing the expected stock price, and hastens to liquidate the shares.

We consider MI functions which are (log-)linear and (log-)quadratic with respect to liquidation speed, and assume noise distributed by the Gamma distribution.

For the noise part of MI, we set $\gamma_n = \gamma$ for $n \in \mathbb{N}$ and

$$\begin{aligned} P(c_l^n - \gamma \in dx) &= \text{Gamma}(\alpha_1/n, n\beta_1)(dx) \\ &= \frac{1}{\Gamma(\alpha_1/n)(n\beta_1)^{\alpha_1/n}} x^{\alpha_1/n-1} e^{-x/(n\beta_1)} 1_{(0,\infty)}(x) dx, \end{aligned}$$

where $\Gamma(x)$ is the Gamma function. Here, α_1, β_1 , and $\gamma > 0$ are constants.

For the deterministic part of MI, we consider two patterns such that $g_n(\psi) = \alpha_0\psi$ and $g_n(\psi) = n\alpha_0\psi^2$ for $\alpha_0 > 0$. In each case, assumptions [A], [B1]–[B3] and [C] are satisfied. The corresponding Lévy measure is

$$\nu(dz) = \frac{\alpha_1}{z} e^{-z/\beta_1} 1_{(0,\infty)}(z) dz.$$

7.1 Log-Linear Impact & Gamma Distribution

Theorem 7 directly implies the following theorem.

Theorem 9. *It holds that*

$$V_t(w, \varphi, s; u_{\text{RN}}) = w + \frac{1 - e^{-\gamma\alpha_0\varphi}}{\gamma\alpha_0} s \quad (7.1)$$

for each $t \in (0, 1]$ and $(w, \varphi, s) \in D$.

The implication of this result is the same as in [6]: the right-hand side of (7.1) is equal to $Ju(w, \varphi, s)$ and converges to $w + \varphi s$ as $\alpha_0 \downarrow 0$ or $\gamma \downarrow 0$, which is the profit gained by choosing the execution strategy of the so-called block liquidation, where a trader sells all shares φ at $t = 0$ when there is no MI. Therefore, the optimal strategy in this case is to liquidate all shares by dividing infinitely within an infinitely short time at $t = 0$ (we refer to such a strategy as a nearly block liquidation at the initial time). Note that the jump part of MI

$$g(\zeta_{r-}) \int_{(0,\infty)} z N(dr, dz)$$

does not influence the value of $V_t(w, \varphi, s; u_{\text{RN}})$.

7.2 Log-Quadratic Impact & Gamma Distribution

In [6], we obtained a partial analytical solution to the problem: when φ is sufficiently small or large, we obtain the explicit form of optimal strategies. However, noise in MI complicates the problem, and deriving the explicit solution is more difficult. Thus, we rely on numerical experiments. Owing to the assumption that the trader is risk-neutral, we may assume that an optimal strategy is deterministic. Here, we introduce the following additional condition.

$$[D] \quad \gamma \geq \frac{\alpha_1 \beta_1}{8}.$$

In fact, we can replace our optimization problem with the deterministic control problem

$$\begin{aligned} f(t, \varphi) &= \sup_{(\zeta_r)_r} \int_0^t \exp \left(- \int_0^r q(\zeta_v) dv \right) \zeta_r dr, \\ q(\zeta) &= \tilde{\mu} + \gamma \alpha_0 \zeta^2 + \alpha_1 \log(\alpha_0 \beta_1 \zeta^2 + 1) \end{aligned}$$

for a deterministic process $(\zeta_r)_r$ under the above assumption. We quote Theorem 5 in [5] below.

Theorem 10. $V_t(w, \varphi, s; u_{\text{RN}}) = w + s f(t, \varphi)$ under [D].

Here, note that Theorems 3–4 imply the following characterization.

Proposition 2. $f(t, \varphi)$ is the viscosity solution of

$$\frac{\partial}{\partial t} f - \sup_{\zeta \geq 0} \left\{ f + \zeta \left(1 - \frac{\partial}{\partial \varphi} f \right) - \gamma \alpha_0 \zeta^2 f + \alpha_1 \log(\alpha_0 \beta_1 \zeta^2 + 1) \right\} = 0 \quad (7.2)$$

with the boundary condition

$$f(0, \varphi) = \varphi, \quad f(t, 0) = 0. \quad (7.3)$$

Moreover, if \tilde{f} is a viscosity solution of (7.2)–(7.3) and has a polynomial growth rate, then $f = \tilde{f}$.

It is difficult to obtain an explicit form of the solution of (7.2)–(7.3). Instead, we solve this problem numerically by considering the deterministic control problem $f_{[nt]}^n(\varphi)$ in the discrete-time model for a sufficiently large n :

$$\begin{aligned} &f_k^n(t, \varphi) \\ &= \sup_{\substack{(\psi_l^n)_{l=0}^{k-1} \subset [0, \varphi]^k, \\ \sum_l \psi_l^n \leq \varphi}} \sum_{l=0}^{k-1} \psi_l^n \exp \left(-\tilde{\mu} \times \frac{l}{n} - \sum_{m=0}^l \left\{ n \gamma \alpha_0 (\psi_l^n)^2 + \frac{\alpha_1}{n} \log(n^2 \alpha_0 \beta_1 (\psi_l^n)^2 + 1) \right\} \right). \end{aligned}$$

We set each parameter as follows: $\alpha_0 = 0.01, t = 1, \tilde{\mu} = 0.05, w = 0, s = 1$ and $n = 500$. For φ , we examine three patterns for $\varphi = 1, 10$, and 100.

7.2.1 The case of fixed γ

In this subsection, we set $\gamma = 1$ to examine the effects of the shape parameter α_1 of the noise in MI. Here, we also set $\beta_1 = 2$. As seen in the numerical experiment in [6], the forms of optimal strategies vary according to the value of φ . Therefore, we summarize our results separately for each φ .

Figure 1 shows the graphs of the optimal strategy $(\zeta_r)_r$ and its corresponding process $(\varphi_r)_r$ of the security holdings in the case of $\varphi = 1$, that is, the number of initial shares of the security is small. As found in [6], if there is no noise in the MI function (i.e., if $\alpha_1 = 0$), then the optimal strategy is to sell up the entire amount at the same speed (note that the roundness at the corner in the left graph of Figure 1 represents the discretization error and is not essential). The same tendency is found in the case of $\alpha_1 = 1$, but in this case the execution time is longer than in the case of $\alpha_1 = 0$. When we take $\alpha_1 = 3$, the situation undergoes a complete change. In this case, the optimal strategy is to increase the execution speed as the time horizon approaches.

When the amount of the security holdings is 10, which is larger than in the case of $\varphi = 1$, the optimal strategy and the corresponding process of the security holdings are as shown in Figure 2. In this case, a trader's optimal strategy is to increase the execution speed as the end of the trading time approaches, which is the same as in the case of $\varphi = 1$ with $\alpha_1 = 3$. Clearly, a larger value of α_1 corresponds to a higher speed of execution closer to the time horizon. We should add that a trader cannot complete the liquidation when $\alpha_1 = 3$: However, as mentioned in Section 5, we can choose a nearly optimal strategy from $\mathcal{A}_1^{\text{SO}}(\varphi)$ without changing the value of the expected proceeds of liquidation by combining the execution strategy in Figure 2 (with $\alpha_1 = 3$) and the terminal (nearly) block liquidation. See Section 5.2 of [6] for details.

When the amount of the security holdings is too large, as in the case of $\varphi = 100$, a trader cannot complete the liquidation regardless of the value of α_1 , as Figure 3 implies. This is similar to the case of $\varphi = 10$ with $\alpha_1 = 3$. The remaining amount of shares of the security at the time horizon is larger for larger noise in MI. Note that the trader can also sell all the shares of the security without decreasing the profit by combining the strategy with the terminal (nearly) block liquidation.

7.2.2 The case of fixed $\tilde{\gamma}$

In the above subsection, we presented a numerical experiment performed to compare the effects of the parameter α_1 by fixing γ . Here, we perform numerical comparison from a different viewpoint.

The results in Section 6 imply that taking the uncertainty in MI into account makes a risk-neutral trader optimistic about the estimation of liquidity risks. To obtain a deeper insight, we investigate the structure of the MI function in more detail. In Theorems 2(ii) and 9, the important parameter is γ , which is the infimum of L_1 and is smaller than (or equal to) $E[L_1]$. We can interpret this as a characteristic feature whereby the (nearly) block liquidation eliminates the effect of positive jumps of $(L_t)_t$. However, there is another decomposition of L_t such that

$$L_t = \tilde{\gamma}t + \int_0^t \int_{(0,\infty)} z \tilde{N}(dr, dz),$$

where

$$\tilde{\gamma} = \gamma + \int_{(0,\infty)} z\nu(dz) = \mathbb{E}[L_1]$$

and

$$\tilde{N}(dr, dz) = N(dr, dz) - \int_{(0,\infty)} z\nu(dz)dr.$$

This representation is essential from the viewpoint of martingale theory. Here, $\tilde{N}(\cdot, \cdot)$ is the compensator of $N(\cdot, \cdot)$ and $\tilde{\gamma}$ can be regarded as the “expectation” of the noise in MI. Just for a risk-neutral world (in which a trader is risk-neutral), as studied in Section 6, we can compare our model with the case of deterministic MI functions as in [6] by setting $\tilde{\gamma} = 1$. Based on this, we conduct another numerical experiment with a constant value of $\tilde{\gamma}$.

Note that in our example

$$\tilde{\gamma} = \gamma + \alpha_1\beta_1 \tag{7.4}$$

and

$$\frac{1}{t}\text{Var}\left(\int_0^t \int_{(0,\infty)} z\tilde{N}(dr, dz)\right) = \alpha_1\beta_1^2 \tag{7.5}$$

hold. Here, (7.4) (respectively, (7.5)) corresponds to the mean (respectively, the variance) of the noise in the MI function at unit time. Comparisons in this subsection are performed with the following assumptions. We set the parameters β_1 and γ to satisfy

$$\gamma + \alpha_1\beta_1 = 1, \quad \alpha_1\beta_1^2 = 0.5.$$

We examine the cases of $\alpha_1 = 0.5$ and 1, and compare them with the case of $\gamma = 1$ and $\alpha_1 = 0$.

Figure 4 shows the case of $\varphi = 1$, where the trader has a small amount of security holdings. Compared with the case in Section 7.2.1, the forms of all optimal strategies are the same; that is, the trader should sell the entire amount at the same speed. The execution times for $\alpha_1 > 0$ are somewhat shorter than for $\alpha_1 = 0$.

Figure 5 corresponds to the case of $\varphi = 10$. The forms of the optimal strategies are similar to the case of $\varphi = 10$, $\alpha_1 = 0, 1$ in Section 7.2.1. Clearly, the speed of execution near the time horizon increases with increasing α_1 .

The results for $\varphi = 100$ are shown in Figure 6. The forms of the optimal strategies are similar to the case of $\varphi = 100$ in Section 7.2.1. However, in contrast to the results in the previous subsection, the remaining amount of shares of the security at the time horizon is smaller for larger α_1 .

Finally, we investigate the total MI cost introduced in [7]:

$$\text{TC}(\Phi_0) = -\log \frac{V_T(0, \Phi_0, s)}{\Phi_0 s}. \tag{7.6}$$

When the market is fully liquid and there is no MI, then the total proceeds of liquidating Φ_0 shares of the security are equal to $\Phi_0 s$, because the drift of the price process is negative and the expected price decreases as time passes (the optimal strategy of a risk-neutral trader is block liquidation at the initial time). On the other hand, in the presence of MI, the (optimal) total

proceeds decrease to $V_T(0, \Phi_0, s) = \Phi_0 s \times \exp(-\text{TC}(\Phi_0))$. Thus, the total MI cost $\text{TC}(\Phi_0)$ denotes the loss rate caused by MI in a risk-neutral world.

Figure 7 shows the total MI costs in the cases of $\Phi_0 = 1$ and 10. Here, we omit the case of $\Phi_0 = 100$ because the amount of shares of the security is too large to complete the liquidation unless otherwise combining terminal block liquidations (which may crash the market). In both cases of $\Phi_0 = 1$ and 10, we find that the total MI cost decreases by increasing α_1 . Since the expected value $\tilde{\gamma}$ of the noise in MI is fixed, an increase in α_1 implies a decrease in γ and β_1 . Risk-neutral traders seem to be sensitive toward the parameter γ rather than toward α_1 , and thus the trader can liquidate the security without concern about the volatility of the noise in MI. Therefore, the total MI cost for $\alpha_1 > 0$ is lower than that for $\alpha_1 = 0$.

8 Concluding Remarks

In this paper, we generalized the framework in [6] and studied an optimal execution problem with random MI. We defined the MI function as a product of an i.i.d. positive random variable and a deterministic function in a discrete-time model. Furthermore, we derived the continuous-time model of an optimization problem as a limit of the discrete-time models, and found out that the noise in MI in the continuous-time model can be described as a Lévy process. Our main results discussed in Sections 3–5 are almost the same as in [6].

When considering uncertainty in MI, there are two typical barometers of the “level” of MI: γ and $\tilde{\gamma}$. By using the former parameter γ , we can decompose MI into a deterministic part $\gamma g(\zeta_t)dt$ and a pure jump part $g(\zeta_{t-}) \int_{(0,\infty)} zN(dt, dz)$. Then, the pure jump part can be regarded as the difference from the deterministic MI case studied in [6]. On the other hand, as mentioned in Sections 6 and 7, the latter parameter $\tilde{\gamma}$ is important not only in martingale theory but also in a risk-neutral world. Studying $\tilde{\gamma}$ also provides some hints about actual trading practices. Regardless of whether we accommodate uncertainty into MI, it may result in an underestimate of MI for a risk-neutral trader. The risk-neutral setting (i.e. the utility function is set as $u_{\text{RN}}(w) = w$) is typical and standard assumption in the study of the execution problem, especially in the limit-order-book model (see, for instance, [1]).

Studying the effects of uncertainty in MI in a risk-averse world is also meaningful. As mentioned in Section 5, when the deterministic part of the MI function is linear, the uncertainty in MI does not significantly influence the trader’s behavior, even when the trader is risk-averse. In future work, we will investigate the case of nonlinear MI.

Finally, in our settings, the MI function is stationary in time, but in the real market, the characteristics of MI change according to the time zone. Therefore, it is meaningful to study the case where the MI function is inhomogeneous in time. This is another topic for future work.

References

- [1] Alfonsi, A., Fruth, A., Schied, A. (2010), “Optimal execution strategies in limit order books with general shape functions,” *Quant. Finance* **10**, 143–157

- [2] Almgren, R. and N. Chriss (2000), “Optimal execution of portfolio transactions,” *J. Risk*, **3**, 5-39.
- [3] Bertsimas, D. and A. W. Lo (1998), “Optimal control of execution costs,” *J. Fin. Markets*, **1**, 1-50.
- [4] Gatheral, J. (2010), “No-dynamic-arbitrage and market impact,” *Quant. Finance* **10**, 749-759
- [5] Ishitani, K. and T. Kato (2009), “Optimal execution problem with random market impact,” *MTEC Journal*, **21**, 83-108. (in Japanese)
- [6] Kato, T. (2012), “Formulation of an optimal execution problem with market impact: derivation from discrete-time models to continuous-time models,” *arXiv preprint*, <http://arxiv.org/pdf/0907.3282>
- [7] Kato, T. (2010), “Formulation of an Optimal Execution Problem with Market Impact,” Proceedings of The 41th ISCIE International Symposium on Stochastic Systems, Theory and Its Applications (SSS’09), 235-240.
- [8] Lions, P.-L. and J.-M. Lasry (2007), “Large investor trading impacts on volatility,” *Paris-Princeton Lectures on Mathematical Finance 2004, Lecture Notes in Mathematics 1919*, Springer, Berlin, 173-190.
- [9] Moazeni, S., Thomas F. Coleman and Y. Li (2011), “Optimal execution under jump models for uncertain price impact,” to appear in *Journal of Computational Finance*.
- [10] Subramanian, A. and R. Jarrow (2001), “The liquidity discount,” *Math. Finance*, **11**, 447-474.
- [11] Nagai, H. (1999), *Stochastic Differential Equations*, Kyoritsu-Shuppan (in Japanese).

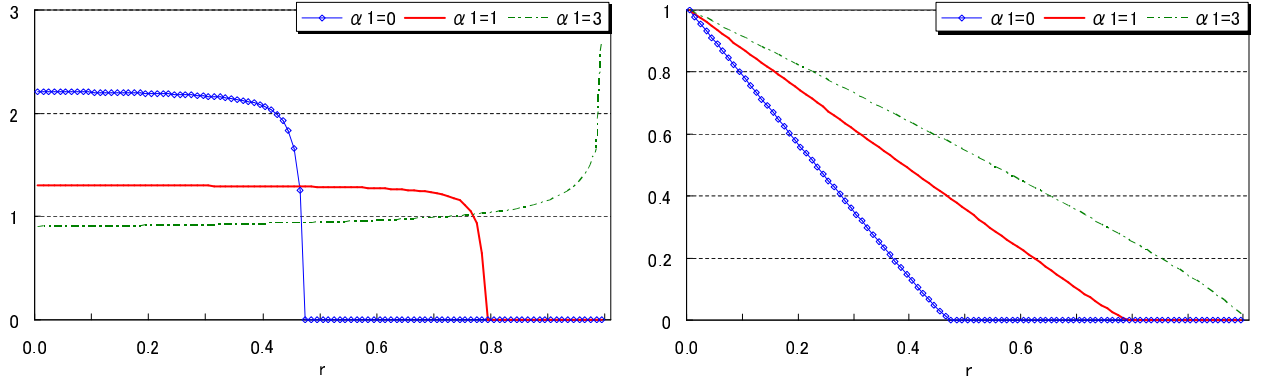


Figure 1: Result for $\varphi = 1$ in the case of fixed γ . Left: The optimal strategy ζ_r . Right: The amount of security holdings φ_r .

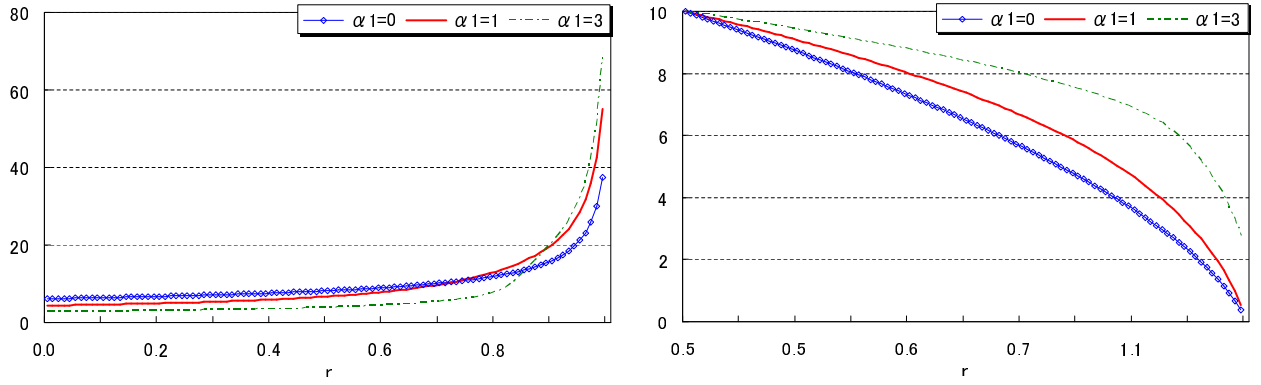


Figure 2: Result for $\varphi = 10$ in the case of fixed γ . Left : The optimal strategy ζ_r . Right : The amount of security holdings φ_r .

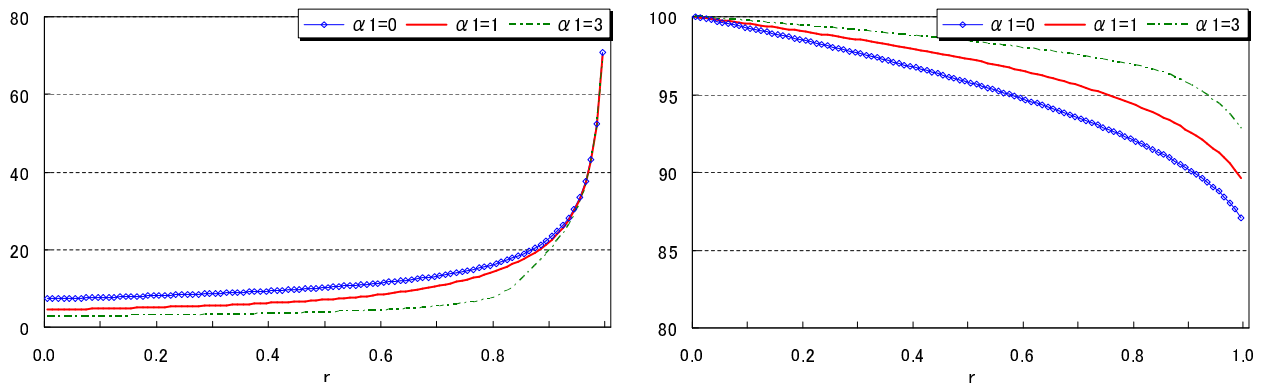


Figure 3: Result for $\varphi = 100$ in the case of fixed γ . Left : The optimal strategy ζ_r . Right : The amount of security holdings φ_r .

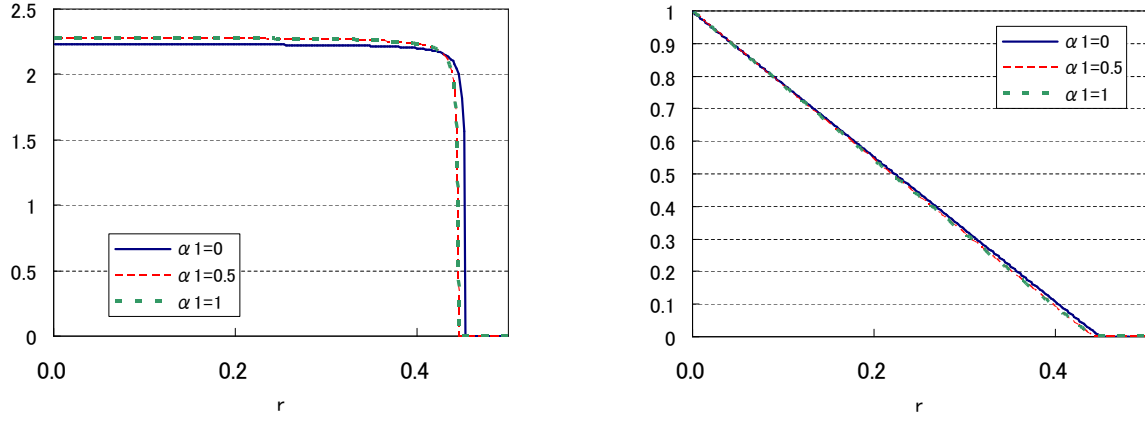


Figure 4: Result for $\varphi = 1$ in the case of fixed $\tilde{\gamma}$. Left : The optimal strategy ζ_r . Right : The amount of security holdings φ_r .

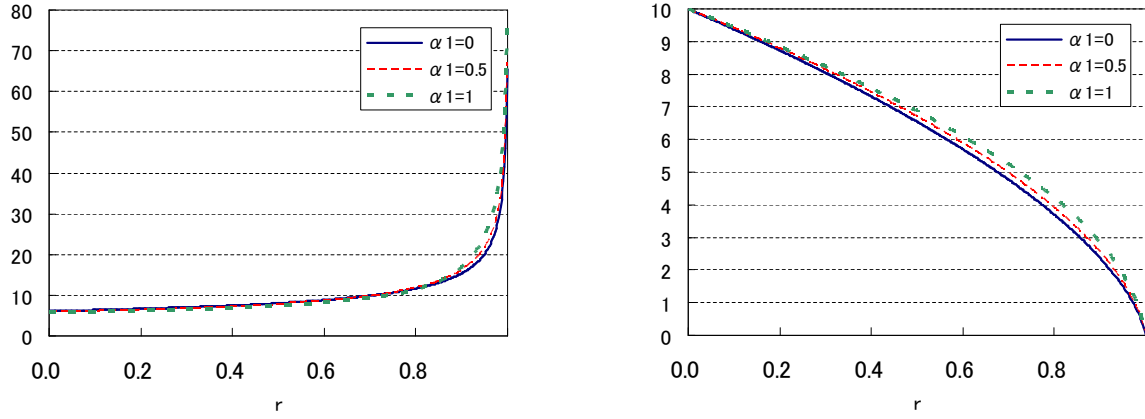


Figure 5: Result for $\varphi = 10$ in the case of fixed $\tilde{\gamma}$. Left: The optimal strategy ζ_r . Right : The amount of security holdings φ_r .

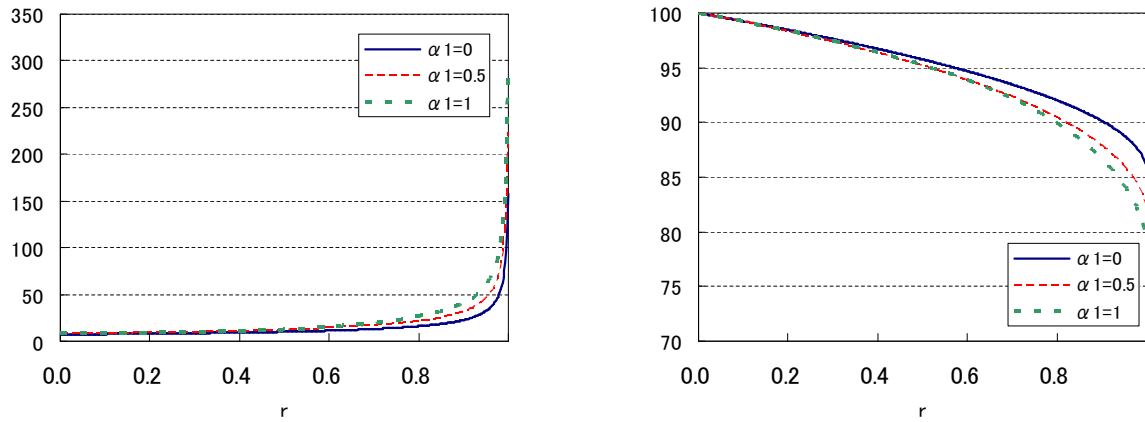


Figure 6: Result for $\varphi = 100$ in the case of fixed $\tilde{\gamma}$. Left : The optimal strategy ζ_r . Right : The amount of security holdings φ_r .

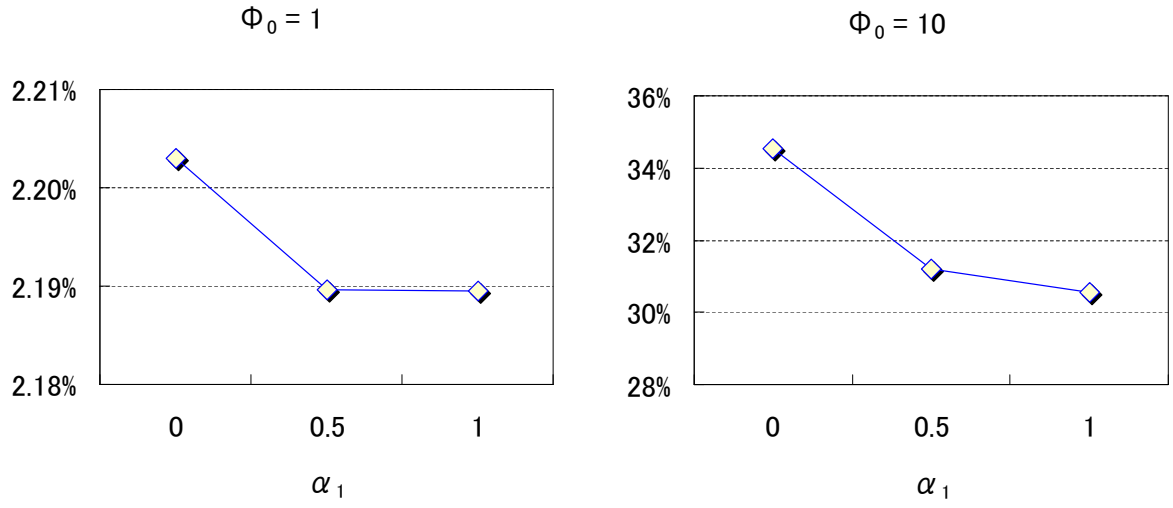


Figure 7: Total MI cost $TC(\Phi_0)$ for a risk-neutral trader. Left: the case of $\Phi_0 = 1$. Right: the case of $\Phi_0 = 10$. The horizontal axes denote the shape parameter α_1 of the Gamma distribution.